“SEMANTICS”

Matt Post

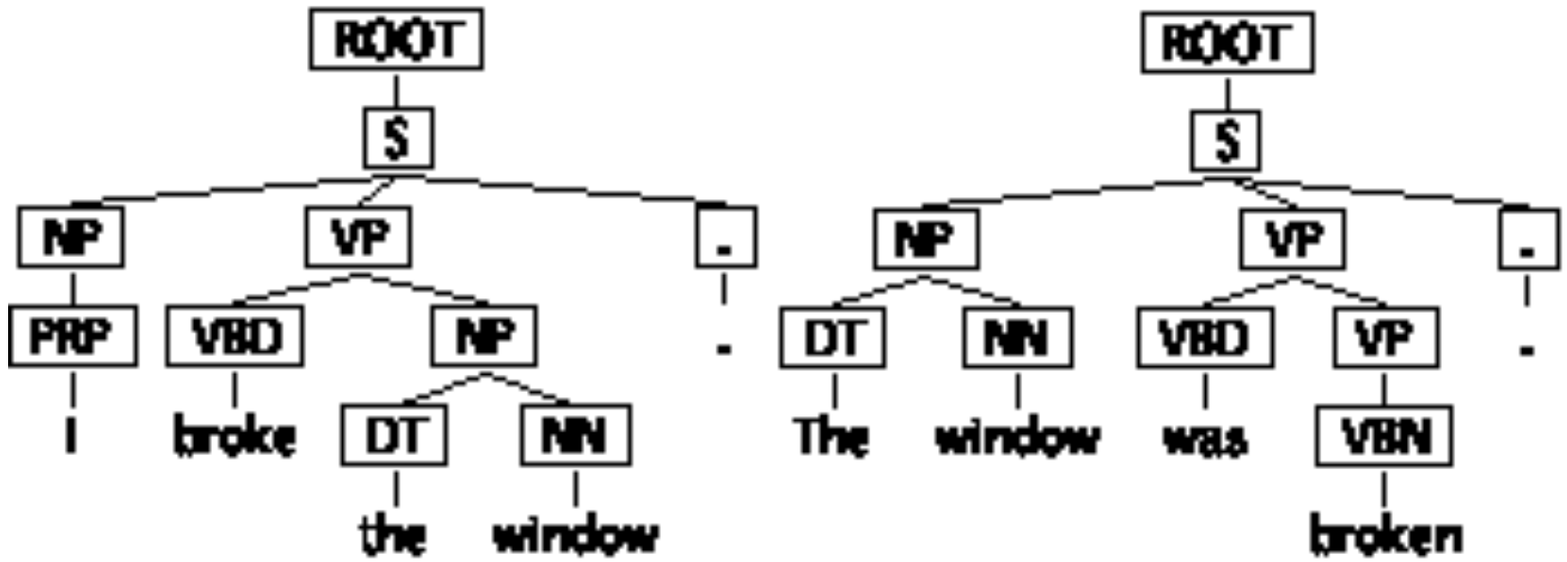
IntroHLT class

23 October 2019

JOHNS HOPKINS
UNIVERSITY

Semantic Roles

- Syntax describes the grammatical relationships between words and phrases
  - But there are many different ways to express a particular meaning



- These variations miss an important generalization

- Structure is important, but one way it is important is as a "scaffolding for meaning"
- What we want to know is

  who *did what to* ***whom***
  and ***when***
  and ***where***
  and ***how****?*

A linguistic hierarchy

Goal

- Given a sentence

Goal

- Given a sentence
  - answer the question “who did what to whom etc”

Goal

- Given a sentence
  - answer the question “who did what to whom etc”
  - store answer in a machine-usable way

Goal

- Given a sentence
  - answer the question “who did what to whom etc”
  - store answer in a machine-usable way
- This requires

Goal

- Given a sentence
  - answer the question “who did what to whom etc”
  - store answer in a machine-usable way
- This requires
  - specifying some representation for meaning

Goal

- Given a sentence
  - answer the question “who did what to whom etc”
  - store answer in a machine-usable way
- This requires
  - specifying some representation for meaning
  - specifying a representation for word relationships

Goal

- Given a sentence
  - answer the question “who did what to whom etc”
  - store answer in a machine-usable way
- This requires
  - specifying some representation for meaning
  - specifying a representation for word relationships
  - mapping the words to these representations

Goal

- Given a sentence
  - answer the question “who did what to whom etc”
  - store answer in a machine-usable way
- This requires
  - specifying some representation for meaning
  - specifying a representation for word relationships
  - mapping the words to these representations
- HOW DO WE REPRESENT MEANING?

Today we will discuss

- An introduction to basic terms of lexical semantics
- WordNet: mapping words to ontologies
- FrameNet: determine the *semantic roles* of words in sentences

Semantics

- What is meaning?
- What is the meaning of the word **cat**?
  - a specific cat?
  - all cats?
  - Platonic ideal of a cat?
  - concept of a cat? (“cat” → CAT)

Much of today’s lecture is borrowed from Philipp Koehn: http://www.inf.ed.ac.uk/teaching/courses/emnlp/

Many meanings

- Example
  - She pays 3% **interest** on the loan.
  - He showed a lot of **interest** in the painting.
  - Microsoft purchased a controlling **interest** in Google.
  - It is in the national **interest** to invade the Bahamas.
  - I only have your best **interest** in mind.
  - Playing chess is one of my **interests**.
  - Business **interests** lobbied for the legislation.
- How many senses is this?

- Another example
  - What is the relationship among these words?
  - {organization, team, group, association, conglomeration, institution, establishment, consortium, federation, agency, coalition, alliance, league, club, confederacy, syndicate, society, corporation}
  - organisation?
- Synonyms often have very different roles
  - {member, part, piece}

Grouping words

- Many-many relationship between form and meaning

Grouping words

- Many-many relationship between form and meaning
- Same forms

Grouping words

- Many-many relationship between form and meaning
- Same forms
  - many related meanings (**polysemy**)

Grouping words

- Many-many relationship between form and meaning
- Same forms
  - many related meanings (**polysemy**)
  - different meanings (**homonymy**)

Grouping words

- Many-many relationship between form and meaning
- Same forms
  - many related meanings (**polysemy**)
  - different meanings (**homonymy**)
- Different forms

Grouping words

- Many-many relationship between form and meaning
- Same forms
  - many related meanings (**polysemy**)
  - different meanings (**homonymy**)
- Different forms
  - same / similar meanings (**synonymy**)

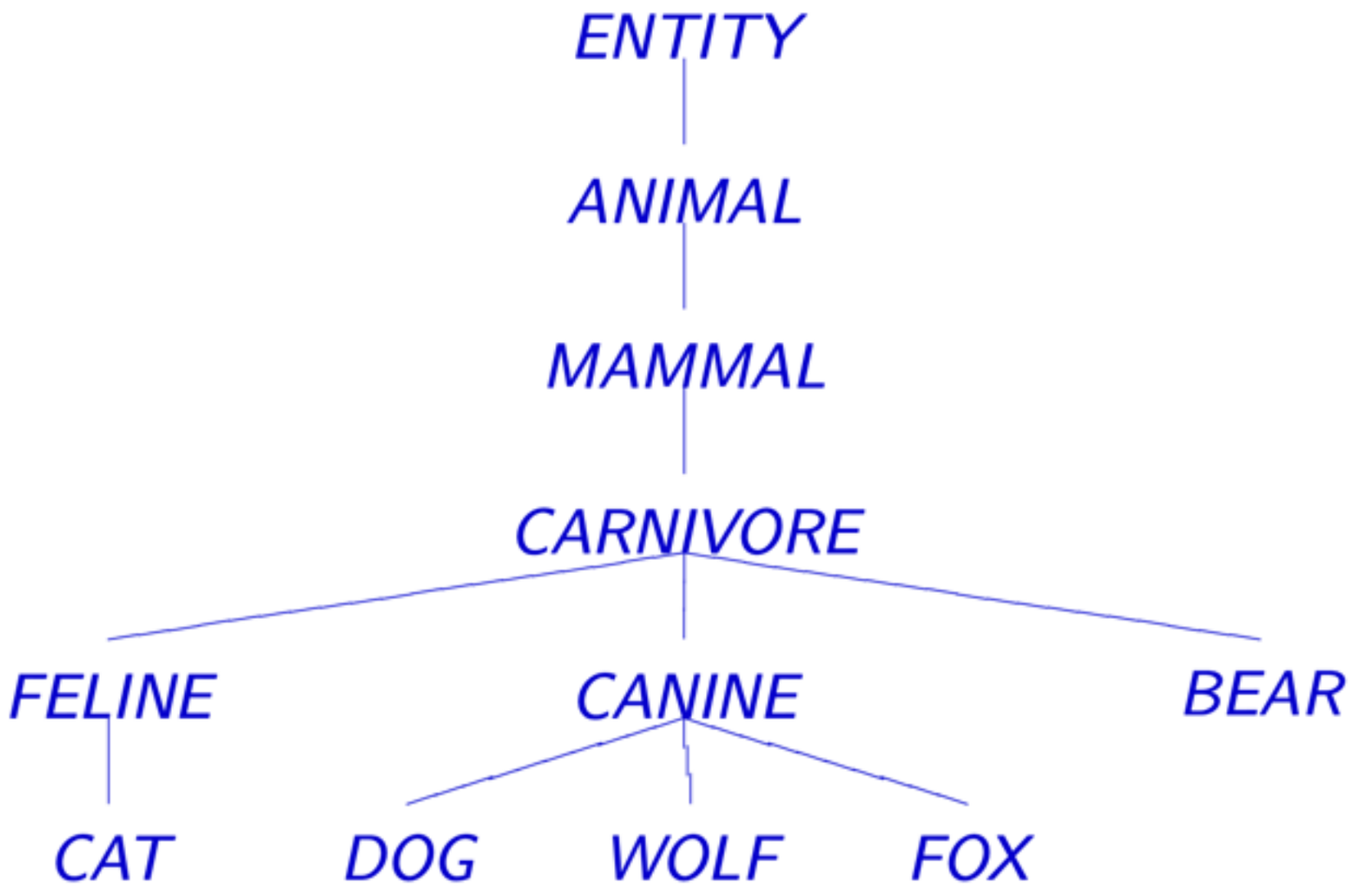
Grouping words

- Many-many relationship between form and meaning
- Same forms
  - many related meanings (**polysemy**)
  - different meanings (**homonymy**)
- Different forms
  - same / similar meanings (**synonymy**)
  - opposite or contrary meaning (**antonymy**)

Relationships among word groups

- Hypernym / hyponym
  - IS-A(animal, cat)
- Part / whole
  - HAS-PART(cat, paw)
  - IS-PART-OF(paw, cat)
- Membership
  - IS-MEMBER-OF(professor, faculty)
  - HAS-MEMBER(faculty, professor)

Ontologies


ENTITY
ANIMAL
MAMMAL
CARNIVORE
FELINE
CANINE
BEAR
CAT
DOG
WOLF
FOX

WordNet

- English WordNet: https://wordnet.princeton.edu/

WordNet Online Demo

WordNet

- English WordNet: https://wordnet.princeton.edu/
- Multilingual WordNet: http://compling.hss.ntu.edu.sg/omw/

WordNet Online Demo

WordNet

- English WordNet: https://wordnet.princeton.edu/
- Multilingual WordNet: http://compling.hss.ntu.edu.sg/omw/
- Words organized into *synsets* (“synonym sets”)

WordNet Online Demo

WordNet Summary

- WordNet represents a particular approach to problem-solving that is reminiscent of earlier symbolic approaches to AI
- The modern approach is more data-driven

Multilingual view of word sense

- Different sense = different translation
- English *interest:*
  - *Zins:* financial charge paid for load (Wordnet sense 4)
  - *Anteil:* stake in a company (Wordnet sense 6)
  - *Interesse:* all other senses
- German *Sicherheit*
  - English *security, safety, confidence*
- English *river*
  - French *fleuve, rivière*

Word Sense Disambiguation

- Back to the **representation** question: how to represent a particular sense of a word?
- Solutions
  - Map to Wordnet or a foreign word sense
  - Map to a *real-life instance of the sense*
- Often depends on the use case
  - search
  - machine translation

WSD as supervised learning problem

- Words can be labeled with their senses
  - *She pays 3%* **interest/INTEREST-MONEY** *on the loan.*
  - *He showed a lot of* **interest/INTEREST-CURIOSITY** *in the painting.*

- Similar to *tagging*
  - given a corpus tagged with senses
  - define features that indicate one sense over another
  - learn a model that predicts the correct sense given the features

- We can apply similar supervised learning methods
  - **Naive Bayes**, related to *HMM*
  - *Transformation-based learning*
  - *Maximum entropy learning*

Simple features

- Directly neighboring words
  - **plant** *life*
  - *manufacturing* **plant**
  - *assembly* **plant**
  - **plant** *closure*
  - **plant** *species*
- Any content words in a 10 word window (also larger windows)
  - *animal*
  - *equipment*
  - *employee*
  - *automatic*

More features

- Syntactically related words
- Syntactic role in sense
- Topic of the text
- Part-of-speech tag, surrounding part-of-speech tags

Training data for supervised WSD

- **SENSEVAL** competition
  - bi-annual competition on WSD
  - provides annotated corpora in many languages

- Pseudo-words
  - create artificial corpus by artificially conflate words
  - example: replace all occurrences of *banana* and *door* with *banana-door*

- Multi-lingual parallel corpora
  - translated texts aligned at the sentence level
  - translation indicates sense

Naive Bayes

- We want to predict the sense S given a set of features F
- First, apply the Bayes rule

$$\text{argmax}_S p(S|F) = \text{argmax}_S p(F|S)p(F) \quad (1)$$

- Then, decompose $p(F)$ by assuming all features are independent (that's *naive*!)

$$p(F) = \prod_{f_i \in F} p(f_i|S) \quad (2)$$

- The *prior* $p(S)$ and the conditional *posterior* probabilities $p(f_i|S)$ can be learned by maximum likelihood estimation

Decision list

- Yarowsky [1994] uses a **decision list** for WSD
  - two senses per word
  - rules of the form: collocation → sense
  - example: *manufacturing* **plant** → *PLANT-FACTORY*
  - rules are ordered, most reliable rules first
  - when classifying a test example, step through the list, make decision on first rule that applies

- Learning: rules are ordered by

$$\log\left(\frac{p(sense_A|collocation_i)}{p(sense_B|collocation_i)}\right) \quad (3)$$

  Smoothing is important

Bootstrapping

- Yarowsky [1995] presents **bootstrapping** method

  1. label a few examples
  2. learn a decision list
  3. apply decision list to unlabeled examples, thus labeling them
  4. add newly labeled examples to training set
  5. go to step 2, until no more examples can be labeled

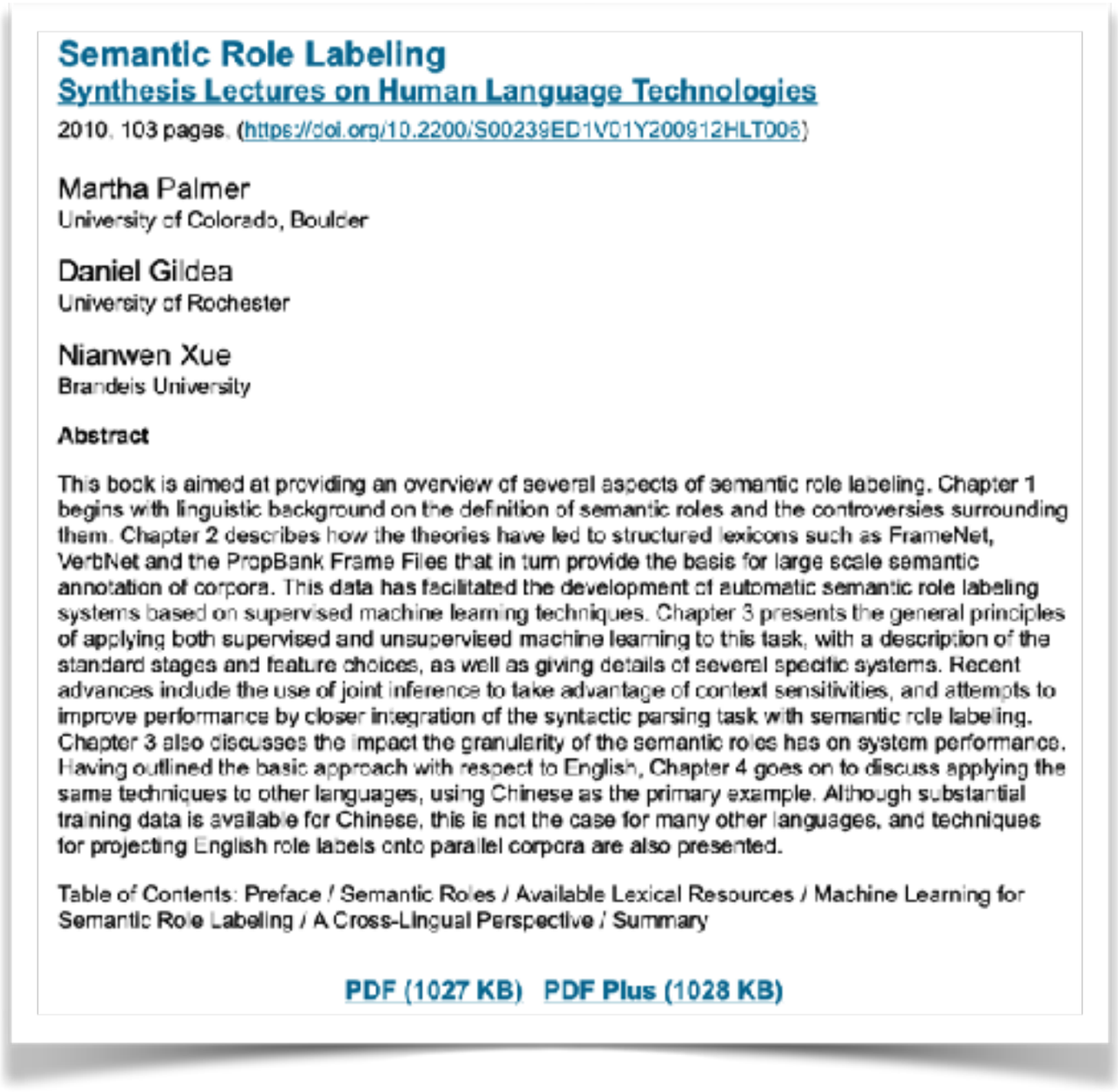
- Initial starting point could also be
  - a short decision list
  - words from dictionary definition

Modern Approaches

- Relies on rich contextualized embeddings of words (e.g., BERT)
- This will be discussed a bit later in the course
  - Information Extraction (Oct. 28)
  - Information Retrieval (Oct. 30)
  - Distributional Semantics (Nov. 4)

Semantic Role Labeling

- Assuming we can disambiguate a word, can we get back to the core question of identifying word relationships?



Semantic Role Labeling
Synthesis Lectures on Human Language Technologies
2010, 103 pages. (https://doi.org/10.2200/S00239ED1V01Y200912HLT006)

Martha Palmer
University of Colorado, Boulder

Daniel Gildea
University of Rochester

Nianwen Xue
Brandeis University

Abstract

This book is aimed at providing an overview of several aspects of semantic role labeling. Chapter 1 begins with linguistic background on the definition of semantic roles and the controversies surrounding them. Chapter 2 describes how the theories have led to structured lexicons such as FrameNet, VerbNet and the PropBank Frame Files that in turn provide the basis for large scale semantic annotation of corpora. This data has facilitated the development of automatic semantic role labeling systems based on supervised machine learning techniques. Chapter 3 presents the general principles of applying both supervised and unsupervised machine learning to this task, with a description of the standard stages and feature choices, as well as giving details of several specific systems. Recent advances include the use of joint inference to take advantage of context sensitivities, and attempts to improve performance by closer integration of the syntactic parsing task with semantic role labeling. Chapter 3 also discusses the impact the granularity of the semantic roles has on system performance. Having outlined the basic approach with respect to English, Chapter 4 goes on to discuss applying the same techniques to other languages, using Chinese as the primary example. Although substantial training data is available for Chinese, this is not the case for many other languages, and techniques for projecting English role labels onto parallel corpora are also presented.

PDF (1027 KB) PDF Plus (1028 KB)

Table 1.1: A set of widely recognized Semantic Roles

Role	Description	Examples
Agent	Initiator of action, capable of volition	**The batter** smashed the pitch into left field. **The pilot** landed the plane as lightly as a feather.
Patient	Affected by action, undergoes change of state	David trimmed **his beard**. John broke **the window**.
Theme	Entity moving, or being "located"	Paola threw **the Frisbee**. **The picture** hangs above the fireplace.
Experiencer	Perceives action but not in control	**He** tasted the delicate flavor of the baby lettuce. **Chris** noticed the cat slip through the partially open door.
Beneficiary	For whose benefit action is performed	He sliced **me** a large chunk of prime rib, and I could hardly wait to sit down to start in on it. The Smiths rented an apartment **for their son**.
Instrument	Intermediary/means used to perform an action	He shot the wounded buffalo with **a rifle**. The surgeon performed the incision with **a scalpel**.
Location	Place of object or action	There are some real monsters hiding **in the anxiety closet**. The band played **on the stage**.
Source	Starting point	The jet took off **from Nairobi**. We heard the rumor **from a friend**.
Goal	Ending point	The ball rolled **to the other end of the hall**. Laura lectured **to the class**.

FrameNet

- One of a handful of resources of this type
- English FrameNet: https://framenet.icsi.berkeley.edu/fndrupal/
- Multilingual FrameNet: https://framenet.icsi.berkeley.edu/fndrupal/node/5549

Task

- Determine semantic roles of words in a sentence
  - Input:

    I saw the bird with the telescope.

  - Output:

    $[\text{I}]_{\text{AGENT}}$ saw $[\text{the bird}]_{\text{THEME}}$ with $[\text{the telescope}]_{\text{INSTR}}$

  - Spans annotated with roles

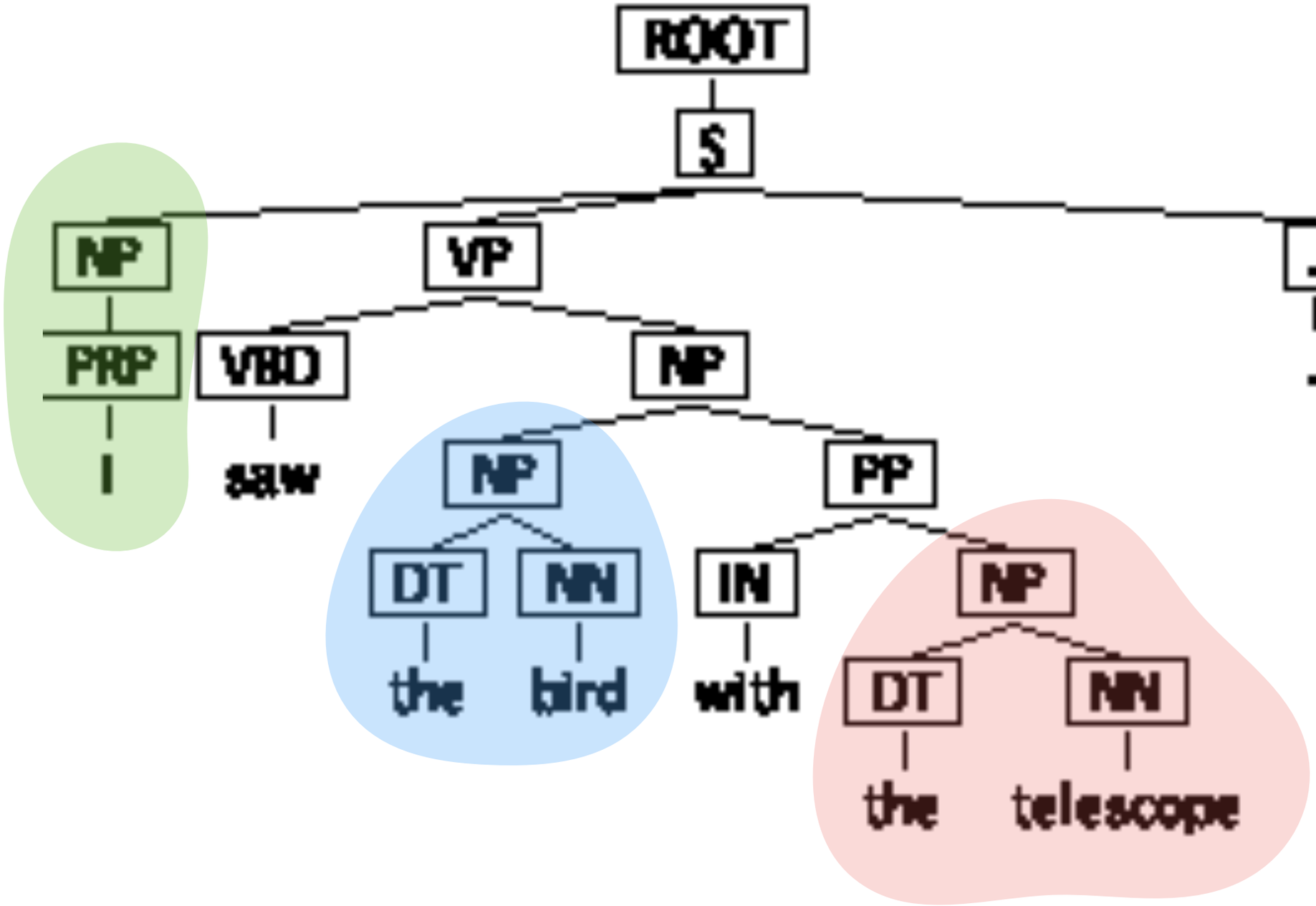
Application

- What kind of task might you apply this to?

Two steps

- Identify the spans
  - Binary decision: is this span a role
- Classification:
  - Categorical decision: what role is it

Observation

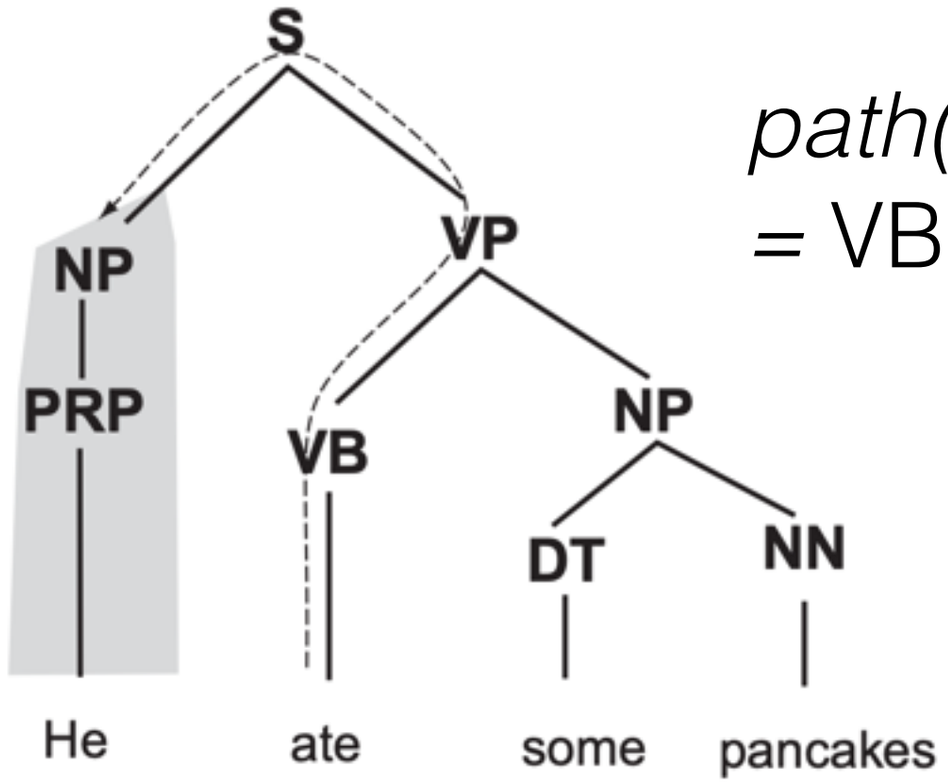


Supervised learning problem

- Training input:
  - span of words
  - correct label
  - host of features
- Training algorithm

Features

- Nonterminal label (“NP”)
- Governing category (“S” or “VP” = subject or object)
- Parse tree path
- Position (before or after predicate)
- Head word
- Many, many more



path(ate→He)
= VB↑VP↑S↓NP

- Trained with discriminative ML algorithms (SVM, MaxEnt)

Other tasks not discussed today

- Sentiment Analysis

There is nothing you can teach a man like Mr. Collard

You could do worse than to buy the Cinetech 12.9 Camera!

→ *positive or negative?*

- Anaphora Resolution

Chris *gave* ***Pat*** *a pat on the back.* ***It*** *didn't help very much.*

Summary

- We have seen an overview of symbolic representations of meaning
  - Extensive statistical and machine learning techniques were used to solve them
  - These dominated research for the past few decades until ~2015
- It is unclear how to represent meaning in a manner that can be operationalized
- Today's end-to-end deep learning approaches have no need for many of these